



Optimal Transport on Completely Integrable Toric Manifolds

Szymon Myga¹ 

Received: 1 August 2019 / Published online: 7 August 2020
© The Author(s) 2020

Abstract

Following the work of Berman and Witt Nyström we present the proof of existence and uniqueness of solutions to transported Monge–Ampère problem on complex compact manifolds with completely integrable torus action. The proof is based on the real theory of optimal transportation and some convex analysis.

Keywords Toric Kahler manifolds · Monge–Ampère equation · Optimal transport

Mathematics Subject Classification 32W20 · 49Q22

1 Introduction

Let (X, ω) be a compact Kähler manifold of real dimension $2n$, i.e., X is a complex manifold and one can find a Hermitian metric on it whose fundamental form ω is closed, thus making (X, ω) into a symplectic manifold. We assume that X is toric—there is a real torus T^k acting on it by automorphisms of ω . Such an action also generates a Lie algebra homomorphism from the Lie algebra of the torus $\mathfrak{t} \simeq \mathbb{R}^k$ into the Lie algebra of the vector fields of X . This action can be extended to the holomorphic action of complexified torus $T_c^k \simeq (\mathbb{C}^*)^k$. We also assume that the action is effective. That is, the trivial automorphism only comes from the identity element. Lastly, we want the action to be Hamiltonian, so we assume that there is a moment map: an action invariant function $m : X \rightarrow (\mathbb{R}^k)^*$, with $(\mathbb{R}^k)^*$ being the Lie algebra dual to \mathfrak{t} , such that for every element $t \in \mathbb{R}^k$

$$-d\langle m(p), t \rangle = \omega_p(t^\#, \cdot)$$

✉ Szymon Myga
szymon.myga@im.uj.edu.pl

¹ Department of Mathematics and Computer Science, Jagiellonian University, Kraków, Poland

with $t^\#$ being the vector field generated by t and $\langle m(p), t \rangle$ being the value of the linear form $m(p)$ at t .

In this setting one can prove that the image of X through m is a compact convex polytope in \mathbb{R}^k with non-empty interior. Moreover this image does not depend on the choice of particular ω in an invariant cohomology class.

Suppose a probability measure with density $1/C < g(p) < C$ is given on a moment polytope P for some toric Kähler manifold (X, ω) . Following the preprint [3] one can define a notion of complex transported Monge–Ampère measure MA_g on X which corresponds to the Monge–Ampère measure that appears in the theory of optimal transportation of measures. Then the natural question to ask is whether the equation

$$MA_g(\phi) = \mu$$

has a unique (up to an additive constant) solution for any invariant measure μ that does not put any mass on polar sets, i.e., the sets that are $-\infty$ loci of plurisubharmonic functions. This is a technical assumption that comes up when one tries to define a Monge–Ampère operator for singular functions. The question of existence was already settled in [3]. The authors also proved there the uniqueness of solutions for the measures of finite pluricomplex energy.

When $k = n$, the action is called completely integrable. This is a special case of torus action, since it is the maximal possible dimension of a torus acting in an effective and Hamiltonian way. From now on we assume complete integrability. In that case the above problem simplifies, thus allowing one to use the full strength of the real theory of optimal transportation and prove uniqueness of solutions in all cases. That was already noticed by Berman and Witt Nyström in [3], although they did not provide the proof. In this note we want to fill this gap and write down the proof of the following

Theorem *For any invariant probability measure μ that does not put mass on polar sets there is an invariant $\phi \in \text{PSH}(X, \omega)$ such that*

$$g(m_\phi)MA(\phi) = \mu,$$

where m_ϕ is a moment map for the torus action induced by ϕ .

The general idea, dating as least to the work of Wang and Zhu [11] is to use the moment map in order to reduce the complex Monge–Ampère equation to the real one. As already suggested in [3] the proof then follows from the result of McCann [7], although not in a straightforward way. The only gap to fill from there is to ensure that appropriate notions of convergence for real and complex solutions coincide. Towards this end we prove the following lemma:

Lemma *If a uniformly Lipschitz sequence of convex functions F_n converges in a monotone way to a convex function F , then their Legendre transforms F_n^* converge to F^* in $W_{\text{loc}}^{1,\infty}$.*

As already mentioned, in the real setting this is a well studied equation that appears in the theory of optimal transportation. In the complex case, it comes up as an equation for Kähler–Ricci solitons on Fano varieties, although in that case the action might not be completely integrable.

2 Background Material

2.1 Convex Functions

Here we want to recall a few facts about convex functions. For a convex function $u : \mathbb{R}^n \rightarrow \mathbb{R} \cup \{+\infty\}$ we define its domain as the convex set $\{x : u(x) < \infty\}$ and denote it as $\text{dom}(u)$. For convenience we exclude the function $u \equiv +\infty$ from the set of convex functions.

For any convex function u on \mathbb{R}^n its Legendre transform is defined by

$$u^*(p) := \sup_{x \in \mathbb{R}^n} \{\langle x, p \rangle - u(x)\}.$$

It is a crucial notion in convex analysis. It is not hard to show that the Legendre transform u^* is a convex lower semicontinuous function.

The multivalued subgradient of u is a set-valued map defined on $\text{int}(\text{dom}(u))$ that attaches to a point the set of slopes of supporting planes at that point, namely

$$y \in \partial u(x) \Leftrightarrow \forall z \in \mathbb{R}^n \quad u(x) + \langle y, z - x \rangle \leq u(z).$$

Since u is convex ∂u is always non-empty on $\text{int}(\text{dom}(u))$. It is single-valued iff u is differentiable at x and at this point it is equal to $\nabla u(x)$.

The notion of a subgradient is closely related to the notion of a Legendre transform through the following equivalences

$$x \cdot p = u(x) + u^*(p) \Leftrightarrow p \in \partial u(x) \Leftrightarrow x \in \partial u^*(p).$$

From this, one can see that in the case of a smooth strictly convex function the gradient of the function and of its Legendre transform are each other's bijective inverses.

The most important fact concerning the differentiability of convex functions is the following one:

Theorem 2.1 (Rademacher's Theorem) *Any convex function is differentiable on a subset of full measure of its domain.*

The set of differentiability of u will be denoted by $\text{dom}(\nabla u)$.

Finally, we list the properties of convex subgradients that will be of use to us.

Proposition 2.2 (Closedness of subgradients, [8, Theorem 24.4]) *For any convex function u the graph of ∂u is a closed subset of $\mathbb{R}^n \times \mathbb{R}^n$.*

Proposition 2.3 (Boundedness of subgradients, [8, Corollary 24.5.1]) *Pick any $x \in \text{int}(\text{dom}(\phi))$, then for any positive ϵ there is a positive δ such that*

$$\partial u(B(x, \delta)) \subseteq \partial u(x) + B(0, \epsilon),$$

with $B(x, \delta)$ denoting the ball centered at x , with the radius δ .

This gives us the following corollary:

Corollary 2.4 *For any convex function u and for any compact subset K of $\text{int}(\text{dom}(u))$ the set $\partial u(K)$ is bounded.*

Proof Indeed, pick an ϵ and take appropriate δ_x -balls at points x in K . This gives a covering of K . The only thing left to show is that at $x \notin \text{dom}(\nabla u)$ the subgradient $\partial u(x)$ is still bounded. But x is in $\text{int}(\text{dom}(u))$, if the subgradient would be unbounded then u would get arbitrarily big arbitrarily close to x , which cannot be since x lies a positive distance away from the boundary of $\text{dom}(u)$. \square

Lemma 2.5 *Let u, v be two convex functions defined on some convex set C with non-empty interior. If $\{\nabla u = \nabla v\}$ is a subset of full measure of C then $u \equiv v$ in $\text{int}(C)$ modulo additive constant.*

Proof If ρ_ϵ is the standard mollifier then it is easy to verify that $u * \rho_\epsilon$ and $v * \rho_\epsilon$ are convex in C^ϵ smooth and converge locally uniformly to u and v , respectively. The set C^ϵ is the domain of definition of mollified function, i.e., $\{x \in C \mid \text{dist}(x, \partial C) > \epsilon\}$. By smoothness, the almost everywhere equality of their gradients implies their equality everywhere and thus $u * \rho_\epsilon$ and $v * \rho_\epsilon$ must converge to the same function up to a constant. \square

2.1.1 Convergence of Convex Functions

The following facts about the convergence of convex functions will be of use. The most natural notion is the following.

Lemma 2.6 *If a sequence of convex functions $\{u_k\}$ converges locally uniformly to a function u , then u is convex.*

We would also like to say something about the convergence of subgradients.

Definition We say that the sequence of subgradients ∂u_n converges graphically to subgradient ∂u if their graphs converge as sets, i.e.,

$$\begin{aligned} \text{graph}(\partial u) = \{(x, p) \mid \exists (x_n, p_n) \in \mathbb{R}^n \times \mathbb{R}^n : p_n \in \partial u_n(x_n) \\ \text{and } (x_n, p_n) \rightarrow (x, p)\}. \end{aligned}$$

The following theorem of Attouch is a fundamental result concerning the graphical convergence of subgradients.

Theorem 2.7 ([9, Theorem 12.35]) *For convex, lower semicontinuous functions f_n and f the following are equivalent:*

1. $f_n \rightarrow f$ locally uniformly,
2. $\partial f_n \rightarrow \partial f$ graphically and for some choice of $p_n \in \partial f_n(x_n)$ and $p \in \partial f(x)$ such that $(x_n, p_n) \rightarrow (x, p)$ one has $f_n(x_n) \rightarrow f(x)$.

Remark The first part of the original theorem is expressed in terms of “epigraphical” convergence in the reference, but it is equivalent to locally uniform convergence (see [9, Theorem 7.17]).

Finally, we would like to describe the relationship between the graphical convergence and pointwise convergence, for this we need the following definition:

Definition The sequence of set-valued maps S_n is equicontinuous at point x with respect to subset X if for each positive ϵ there is a neighbourhood V of x such that for almost every n

$$S_n(y) \subset S_n(x) + B(0, \epsilon) \text{ for all } y \in V \cap X.$$

The sequence is equicontinuous with respect to X if it is equicontinuous at each point of X .

Remark In [9] the above notion is called the asymptotic equi-outer-semicontinuity. Since we don’t need other notions of equicontinuity we will just call that one the equicontinuity.

Now the desired relationship is the following:

Theorem 2.8 ([9, Theorem 5.40]) *For the sequence of set-valued maps S_n , the map S and a set X any pair of the following conditions implies the third:*

1. S_n is equicontinuous with respect to X ,
2. S_n converges graphically to S relative to X ,
3. S_n converges pointwise to S relative to X .

2.1.2 The Class of Globally Lipschitz Convex Functions

Definition By support function of a bounded convex set P containing zero we mean the function

$$\phi_P(x) := \sup_{p \in P} \langle x, p \rangle.$$

If the set P is bounded then ϕ_P is finite everywhere.

Definition We will denote by \mathcal{P} the space of convex functions dominated by ϕ_P , i.e., the set $\{u - \text{convex} \mid \exists C : u \leq \phi_P + C\}$. This is the set of convex functions whose Legendre transform is $+\infty$ outside P . The subset of \mathcal{P} consisting of functions that also dominate ϕ_P will be denoted by \mathcal{P}_+ , in other words $\mathcal{P}_+ = \{u \in \mathcal{P} \mid \exists C : u - C \geq \phi_P\}$. Both sets can be equipped with the topology of pointwise convergence, which is equivalent to the topology of locally uniform convergence, by the virtue of uniform Lipschitz constant for all \mathcal{P} .

The set \mathcal{P}_+ is dense in \mathcal{P} . Moreover the approximating sequence can be chosen as nice as possible.

Lemma 2.9 ([2, Lemma 2.2]) *Every $\phi \in \mathcal{P}$ can be approximated by decreasing sequence of smooth strictly convex functions from \mathcal{P}_+ .*

2.2 Optimal Transport and Monge–Ampère Equation

We say that the function $T : \mathbb{R}^n \rightarrow \mathbb{R}^n$ transports probability measure μ to probability measure ν if for any Borel set A the following equality holds

$$\nu[A] = \mu[T^{-1}(A)].$$

Alternatively we say that T pushes μ forward to ν and denote the push-forward measure by $T_{\#}\mu$.

In general there will be a lot of such maps, so it is natural to put some optimality constraints on them. The best understood constraint and in some cases the natural one is minimizing the quadratic cost, i.e., the transport map should minimize the following functional

$$\int_{\mathbb{R}^d} |x - T(x)|^2 d\mu.$$

In general there might not be a solution and if it exists it might not be unique, some regularity assumptions for the measures must be added. For example, one can assume that the measures have finite second moments and μ is absolutely continuous. In that case the solution exists and has a form of $T = \nabla\phi$ for some convex function ϕ . For thorough discussion of this problem, the reader might consult [10].

Supposing that a solution exists, by the transport condition we get

$$\int \chi_A d\nu = \int_A d\nu = \int_{(\nabla\phi)^{-1}(A)} d\mu = \int \chi_A \circ \nabla\phi d\mu$$

That can easily be generalized to get that for any $f \in C_b(\mathbb{R}^n)$

$$\int f d\nu = \int f \circ \nabla\phi d\mu. \quad (1)$$

Here $C_b(\mathbb{R}^n)$ denotes the set of continuous and bounded functions on \mathbb{R}^n .

Suppose now that $d\nu = g(x)dx$ for some density $g(x)$ and ϕ is a C^2 function. By change of variables formula we get that

$$\int f(\nabla\phi(x)) d\mu = \int f(\nabla\phi(x)) g(\nabla\phi(x)) \det D^2\phi dx$$

and that provides one with a notion of solution to the transported Monge–Ampère equation

$$MA_g^{\mathbb{R}}(\phi) := g(\nabla\phi(x)) \det D^2\phi = \mu$$

as long as the optimal transport map exists.

As we mentioned, for any two probability measures the optimal transport solution might not exist. However, under a mild regularity assumption it is still possible to transport one to another through a subgradient of convex function, so that the condition (1) is still satisfied. This is the content of the following important theorem.

Theorem 2.10 (McCann [7]) *Let μ, ν be probability measures on \mathbb{R}^n and suppose that μ vanishes on Borel subsets of \mathbb{R}^n of Hausdorff dimension $n - 1$. Then there exists a convex function ψ on \mathbb{R}^n whose subgradient $\partial\psi$ pushes μ forward to ν . $\partial\psi$ is uniquely determined μ -almost everywhere.*

Of course the assumption on the null sets of μ can not be abandoned. For example if $\mu = \delta_x$ and ν is not a point measure, then if A is such a set that $0 < \nu[A] < 1$ one gets that for any convex function ϕ , $\nu[A] \neq \mu[(\partial\phi)^{-1}(A)]$ since the latter must always be either 0 or 1.

2.3 Torus Action

As Sect. 1 we are interested in completely integrable Kähler manifolds. In this setting the following results provide the correspondence between the Kähler geometry and convex functions.

Proposition 2.11 ([5]) *There is an open dense subset $X_0 \subset X$ where the action of T_c^n is free, making X_0 diffeomorphic to $(\mathbb{C}^*)^n$. Every invariant Kähler form ω on X has a Kähler potential on X_0 , i.e.,*

$$\omega|_{X_0} = 2i\partial\bar{\partial}F$$

for some F .

The set $X \setminus X_0$ is given as a vanishing set of some holomorphic vector fields, so it must be analytic.

If we introduce coordinates on X_0 coming from $(\mathbb{C}^*)^n$ by $L : e^{x+iy} \rightarrow x + iy$, the invariance of potential means that the function F from the previous proposition depends only on x variable in \mathbb{R}^n and positive definiteness means that F must be convex. Moreover, nothing in the proof actually requires the form to be smooth, so the conclusion easily extends to forms with more singular coefficients, thus asserting that every closed positive and invariant $(1, 1)$ -current in the cohomology class $[\omega]$ admits a convex potential.

Proposition 2.12 *For the symplectic form ω as above, the moment map is*

$$\frac{\partial F}{\partial x} + c,$$

with c being any constant vector in \mathbb{R}^n .

Finally, we recall the theorem of Atiyah [1], Guillemin and Sternberg [6]:

Theorem 2.13 *The image of X through the moment map is a compact convex polytope in \mathbb{R}^n .*

In the case of completely integrable actions the polytopes that can arise as images of moment maps are called Delzant polytopes. Conversely for each Delzant polytope there exists a Kähler manifold with completely integrable torus action and a moment map that maps to this polytope.

2.4 Toric Pluripotential Theory

The class of plurisubharmonic functions that are torus invariant will be denoted by $PSH_{\text{tor}}(X, \omega) = \{\phi \in PSH(X, \omega) \mid \forall z \in X, t \in T^n \mid \phi(t \cdot z) = \phi(z)\}$. The results of the previous section imply that to each such function corresponds a convex function on \mathbb{R}^n .

More precisely, if set X_0 are coordinate map L are as in the previous subsection then for any $v \in PSH_{\text{tor}}(X, \omega)$ the form $\omega_v = \omega + i\partial\bar{\partial}v$ is still invariant and closed there, so restricting to X_0 there is a convex F_v function given by

$$F_v \circ L = F_0 \circ L + v,$$

with $F_0 \circ L$ being the potential for ω . Of course if use the formula above to produce a plurisubharmonic function it will only be defined on X_0 , but since $X \setminus X_0$ is analytic, the function will extend to the whole X .

Not every convex function can be a potential for an invariant Kähler form. If P is the Delzant polytope of the manifold (X, ω) then the following Proposition holds (see, e.g., [4] for a proof).

Proposition 2.14 *The following are equivalent:*

1. $v \in PSH_{\text{tor}}(X)$,
2. $F_v \in \mathcal{P}$.

Finally, the two Monge–Ampère measures coincide up to a constant. Specifically, the complex Monge–Ampère measure is defined as

$$MA^{\mathbb{C}}(v)[A] := \int_A (\omega + i\partial\bar{\partial}v)^n$$

and the real measure as

$$MA^{\mathbb{R}}(F)[B] := |\partial F(B)|$$

for Borel sets A, B in \mathbb{C}^n and \mathbb{R}^n , respectively. Here $|\cdot|$ is the Lebesgue measure. For C^2 convex functions it coincides with the measure

$$MA^{\mathbb{R}}(F)[B] := \int_B \det(D^2 F).$$

For invariant plurisubharmonic function the two concepts are connected through the following proposition.

Proposition 2.15 *Let $\phi \in PSH_{\text{tor}}(X)$, we identify X_0 with $(\mathbb{C}^*)^n$. Then for any $f \in C_b(\mathbb{R}^n)$*

$$\int_{X_0} (f \circ L) MA^{\mathbb{C}}(\phi) = \frac{n!}{(2\pi)^n} \int_{\mathbb{R}^n} f MA^{\mathbb{R}}(F_{\phi}).$$

The proof is just a straightforward computation (see, e.g., [4]) in the smooth case and then the application of classical convergence theorems for convex and plurisubharmonic functions.

2.4.1 g-Monge–Ampère Measure

Following the preprint [3] we define the complex g -Monge–Ampère measure or the complex transported Monge–Ampère measure as

$$MA_g(\phi) := g(m_{\phi})MA(\phi).$$

From the Proposition 2.12 and the definition of the real transported Monge–Ampère measure it is not hard to see that 2.15 extends for smooth functions to transported measures. In the more general case, especially with the torus of smaller rank, the definition becomes more intricate.

If we denote by \mathcal{E}_g the set of all PSH_{tor} functions with full MA_g mass, i.e., those functions for which

$$\int_X MA_g(\phi) = \int_P g(p)dp = 1$$

then the following crucial continuity statement holds:

Theorem 2.16 ([3], Theorem 2.7) *If ϕ_j is a sequence in \mathcal{E}_g decreasing to ϕ in \mathcal{E}_g then*

$$MA_g(\phi_j) \rightarrow MA_g(\phi)$$

in the weak topology of measures.

3 Full Rank Existence and Uniqueness

Given a probability measure $g(p)dp$ on P and any probability measure μ on \mathbb{R}^n , we would like to solve the equation

$$MA_g^{\mathbb{R}}(u) = \mu$$

in some appropriate sense. One can not apply McCann's theorem directly since for example $\mu = \delta_x$ would prevent the existence of the transport map, thus we must use the regularity of g .

Suppose that we have a smooth strictly convex solution u , so that every term in $MA_g^{\mathbb{R}}(u)$ is well-defined and moreover so is ∇u^* . By the fact that for any x and any p , $\nabla u(\nabla u^*(p)) = p$ and $\nabla u^*(\nabla u(x)) = x$ we define the solution through the change of variables formula. Thus

$$\int_{\mathbb{R}^n} f(x)g(\nabla u(x))MA^{\mathbb{R}}(u) = \int_P f(\nabla u^*(p))g(p)dp = \int_{\mathbb{R}^n} f d\mu \quad (2)$$

and a function $u \in \mathcal{P}$ such that the second equality holds for any continuous bounded function f is defined to be a solution.

The fact that there is such a solution follows easily from McCann's theorem. Suppose that ϕ is the convex function whose gradient transports $g(p)dp$ (understood as a measure on \mathbb{R}^n) to μ . By the regularity of g and McCann's theorem it must exist. Then $\nabla \phi$ is defined $g(p)dp$ -almost everywhere and since P is convex we can take ϕ to be $+\infty$ outside of P . Thus after possibly fixing ϕ on ∂P so that it is lower semicontinuous, its Legendre transform ϕ^* becomes unique and defined everywhere on \mathbb{R}^n and thus belongs to the class \mathcal{P} since by lower semicontinuity $\phi^{**} = \phi$. The convex function $u = \phi^*$ is the unique (up to additive constant) solution to the transported Monge–Ampère problem in the class \mathcal{P} . Indeed, since $\nabla \phi$ transports $g(p)dp$ to μ it means that for any $f \in C_b(\mathbb{R}^n)$

$$\int_{\mathbb{R}^n} f d\mu = \int_P f(\nabla \phi(p))g(p)dp = \int_P f(\nabla u^*(p))g(p)dp.$$

If there was to be another solution v in the class \mathcal{P} then its Legendre transform would have been $+\infty$ on the complement of P and lower semicontinuous on its boundary and it would induce a transport of $g(p)dp$ to μ , so by McCann's uniqueness theorem $\nabla u^* = \nabla v^*$ $g(p)dp$ -almost everywhere, and since $g > 0$, by Lemma 2.5 we get $u^* = v^*$ (mod \mathbb{R}) everywhere on $\text{int}(P)$.

Remark on Uniqueness Of course the uniqueness statement becomes false if we allow functions outside of class \mathcal{P} . Suppose that $\mu = \delta_0$, then the solution is obviously $u = \phi_P$, so that $u^* \equiv 0$ on P . But now adding to u any convex function v such that $\min v = v(0)$ would also give a solution, since $(u + v)^* \equiv 0$ on P .

Example Finally, we would like to point out that if the assumptions of McCann's theorem are not satisfied for at least one of the measures there might not be a weak solution. For example, if $\mu = \delta_{[-1,1] \times \{0\}}$ and $\nu = \delta_{\{0\} \times [-1,1]}$ are measures on \mathbb{R}^2 , then it is impossible to find a solution. Indeed, the only candidate is $u(x, y) = |y|$ and it is easy to see that it can not be the solution since $u^*|_{\{0\} \times [-1,1]} \equiv 0$.

3.1 The Complex Case

Corollary 3.1 *The solution to the real problem in \mathbb{R}^n induces a unique solution to the g -Monge–Ampère problem on toric manifolds.*

Proof Firstly, we notice that the fact that μ does not put any mass on pluripolar sets implies that $X \setminus X_0$ as an analytic set has no mass. Thus we can restrict the problem to X_0 . Moreover, since the measure is invariant, it can be interpreted as a measure on \mathbb{R}^n also denoted by μ .

Suppose now we have a real solution F_ϕ for the measure μ , then one suspects that $\phi = (F_\phi - F_0) \circ L$ would be the solution for the corresponding invariant measure. Indeed, F_ϕ is in \mathcal{P} , so it must correspond to some invariant psh function. Moreover, for smooth strictly convex functions the formula (2) obviously translates by Proposition 2.15 to the complex setting. Finally, by Lemma 2.9 there exists a decreasing sequence F_n of smooth strictly convex functions that decreases to F_ϕ , so by smoothness and Theorem 2.16 $MA_g^{\mathbb{R}}(F_n) = MA_g^{\mathbb{C}}(F_n - F_0)$ converges weakly to $MA_g^{\mathbb{C}}(F_\phi)$. Thus the only thing left to show is that $MA_g^{\mathbb{R}}(F_n)$ converges weakly to μ .

Take $f \in C_b(\mathbb{R}^n)$ and put $f_n := f \circ \nabla F_n^*$. We would like to show that f_n converge almost everywhere to f . That would give us the desired assertion by the dominated convergence.

First, let us prove that decreasing convergence of F_n implies locally uniform convergence of F_n^* . Indeed, since F_n 's are uniformly Lipschitz, their pointwise convergence implies locally uniform convergence. Now $F \leq F_n$ implies $F_n^* \leq F^*$, take $p \in \text{int} P$ and suppose that the supremum in $F^*(p)$ is realized by x^* , thus

$$F^*(p) - F_n^*(p) = \sup_{x \in \mathbb{R}^n} \{\langle x, p \rangle - F(x)\} + \inf_{x \in \mathbb{R}^n} \{F_n(x) - \langle x, p \rangle\} \leq F_n(x^*) - F(x^*).$$

Thus F_n^* converges pointwise to F^* . If $K \subseteq \text{int}(P)$ is compact then by Proposition 2.2 and Corollary 2.4 $\partial F^*(K)$ is compact and for every $q \in K$ the supremum in $F^*(q)$ is realized by some y^* in $\partial F^*(q)$, thus the convergence is locally uniform.

Now, we will show that F_n^* converging locally uniformly to F^* implies that ∇F_n^* converges to ∇F^* almost everywhere and that would finish the proof. To do that we want to employ Theorems 2.7 and 2.8 restricted to $\text{dom}(\nabla F^*)$. Thus the only thing left to show is the equicontinuity of ∇F_n^* 's with respect to $\text{dom}(\nabla F^*)$. \square

In order to prove this we will first prove the following lemma:

Lemma 3.2 ∇F_n^* are locally bounded independently of n .

Proof Take any point $x \in \text{int} P \cap \text{dom}(\nabla F^*)$ and pick a positive δ such that $B(x, \delta)$ is relatively compact in $\text{int} P$. By the boundedness of the subgradient (Corollary 2.4) there exists a positive M such that

$$\partial F^*(B(x, \delta)) \subseteq B(\nabla F^*(x), M).$$

Now pick a positive η , starting from some n we get that $0 \leq F^* - F_n^* \leq \eta$ over $B(x, \delta)$. We claim that $\nabla F_n^*(B(x, \delta/2)) \subset B(\nabla F^*(x), M + C)$ holds for some constant C , independent of F_n^* .

Indeed, by convexity it is enough to estimate the gradients on the boundary of $B(x, \delta/2)$. Take a point $y \in \partial B(x, \delta/2)$ such that $|\nabla F_n^*(y)|$ achieves maximum over $\partial B(x, \delta/2)$. The vector $\nabla F_n^*(y)$ must be pointed to the outside of $B(x, \delta/2)$ or at least be tangent to it. The “boundary” steepest case is $F_n^*(y) = F^*(y) - \eta$, F^* growing at best possible rate from y and tangent plane at $F_n^*(y)$ touching F^* at the boundary of $B(x, \delta)$, then $\nabla F_n^*(y)$ would become the steepest if that happened over shortest possible interval which would be of length $\delta/2$. Thus finally $|\nabla F_n^*(y)| \leq \eta + p\delta/2$, where p is the length of the longest vector in $B(\nabla F^*(x), M)$. \square

With the Lemma in hand the rest of the proof is straightforward. Suppose the sequence is not equicontinuous at some point $x_0 \in \text{int} P \cap \text{dom}(\nabla F^*)$. Thus there is a positive ϵ such that for any k there is $y_k \in B(x_0, 1/k) \cap \text{dom}(\nabla F^*)$ such that

$$|\nabla F_{n(k)}(x_0) - \nabla F_{n(k)}(y_k)| > \epsilon$$

with $n(k)$ being some subsequence of \mathbb{N} . But by above lemma the set $p_k = \{\nabla F_{n(k)}^*(y_k)\}$ is bounded, thus there must be a convergent subsequence, conveniently also named p_k , such that $p_k \xrightarrow{k \rightarrow \infty} p$. But the set $q_k = \{\nabla F_{n(k)}^*(x_0)\}$ is also bounded thus a subsequence must converge to some q such that $|q - p| \geq \epsilon$. Thus we have two subsequences (y_k, p_k) and (x_0, q_k) . By graphical convergence both of them must converge to some point in $\partial F^*(x_0)$, but this set is a singleton and that is a contradiction. \square

Acknowledgements The author would like to thank Sławomir Dinew for his guidance. The author was supported by Polish National Science Centre grant 2018/29/N/ST1/02817.

Open Access This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by/4.0/>.

References

- Atiyah, M.F.: Convexity and commuting Hamiltonians. *Bull. Lond. Math. Soc.* **14**, 1–15 (1982)
- Berman, R.J., Berndtsson, B.: Real Monge—Ampère equations and Kähler—Ricci solitons on toric log Fano varieties. *Ann. Fac. Sci. Toulouse Math.* **6**(22), 649–711 (2013)
- Berman, R.J., Witt Nyström, D.: Complex optimal transport and the pluripotential theory of Kähler—Ricci solitons. Preprint. [arXiv:1401.8264](https://arxiv.org/abs/1401.8264)
- Coman, D., Guedj, V., Sahin, S., Zeriahi, A.: Toric pluripotential theory. *Ann. Polon. Math.* <https://doi.org/10.4064/ap180409-3-7>
- Guillemin, V.: Kaehler structures on toric varieties. *J. Differ. Geom.* **40**, 285–309 (1994)
- Guillemin, V., Sternberg, S.: Convexity properties of the moment mapping. *Invent. Math.* **67**, 491–513 (1982)
- McCann, R.J.: Existence and uniqueness of monotone measure-preserving maps. *Duke Math. J.* **80**, 309–323 (1995)

8. Rockafellar, R.T.: *Convex Analysis*. Princeton Mathematical Series 28. Princeton University Press, Princeton, NJ (1970)
9. Rockafellar, R.T., Wets, R.J.-B.: *Variational Analysis*. Grundlehren der Mathematischen Wissenschaften 317. Springer, Berlin (1998)
10. Villani, C.: *Optimal Transport. Old and New*. Grundlehren der Mathematischen Wissenschaften 338. Springer, Berlin (2009)
11. Wang, X.-J., Zhu, X.: Kähler–Ricci solitons on toric manifolds with positive first Chern class. *Adv. Math.* **188**, 87–103 (2004)

Publisher's Note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.